AFRL-AFOSR-VA-TR-2016-0289

Discovery of Empirical Components by Information Theory

Amit Singer
TRUSTEES OF PRINCETON UNIVERSITY
1 NASSAU HALL
PRINCETON, NJ 08544-0001

08/10/2016
Final Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/RTA2

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
10-08-2016	Final Performance	15 Feb 2013 to 14 Feb 2016

4. TITLE AND SUBTITLE
Discovery of Empirical Components by Information Theory

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-13-1-0076

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Amit Singer

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
TRUSTEES OF PRINCETON UNIVERSITY
1 NASSAU HALL
PRINCETON, NJ 08544-0001 US

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AF Office of Scientific Research
875 N. Randolph St. Room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/AFOSR RTA2

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-AFOSR-VA-TR-2016-0289

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRIBUTION A: Distribution approved for public release.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Recent years have seen exciting new developments in mathematics and computer science, which have opened up new domains of application for computational mathematics. These come with new challenges, for which new approaches and tools must be and are being developed. Machine learning and compressive sensing are two typical examples; they draw not only from traditional linear algebra based numerical analysis
or approximation theory, but also from information theory, graph theory, the geometry of Banach spaces, probability theory, and more.
This proposal seeks to fund the research of three faculty drawn to these new computational challenges, who are also finding increasingly that their different fields of expertise all contribute to the development of dramatically more effective tools. This confluence of interests, and the conviction that joining their efforts will produce a whole that exceeds the sum of its parts, constitute the engine that drives the approaches proposed here.

15. SUBJECT TERMS
Random Matrix Theory, information theory

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			RIECKEN, Richard
					19b.

Unclassified	Unclassified	Unclassified	UU	**PAGES**	**19b. TELEPHONE NUMBER** *(Include area code)* 703-941-1100

AFOSR Final Performance Report

Project Title: Discovery of Empirical Components by Information Theory

Award Number: FA9550-13-1-0076

Program Officer:
Dr. Tristan Nguyen
Air Force Defense Research Sciences Program
Air Force Office of Scientific Research
875 North Randolph Street, Room 3112
Arlington, VA 22203-1768
E-mail: Tristan.nguyen@us.af.mil
Phone: (703) 696-7796
Fax: (703) 696-7360

Principal Investigator:
Professor Amit Singer
Department of Mathematics and the
Program in Applied and Computational Mathematics
Princeton University
202 Fine Hall
Princeton, NJ 08544
E-mail: amits@math.princeton.edu
Phone: (609) 258-3682
Fax: (609) 258-1735

Subawardees:
Professor Robert Calderbank
Department of Mathematics, Department of Computer Science
and Department of Electrical and Computer Engineering
Duke University
140 Science Drive
317 Gross Hall
Durham, NC 27708
E-mail: robert.calderbank@duke.edu
Phone: (919) 613-7874
Fax: (919) 660-6519

Professor Ingrid Daubechies
Department of Mathematics and Department of
Electrical and Computer Engineering
Duke University
120 Science Drive
Rm 117 Physics Bldg.
Durham, NC 27708
E-Mail: ingrid.daubechies@duke.edu
Phone: (919) 660-2805
Fax: (919) 660-2821

Final Performance Report

To:	Technicalreports@afosr.af.mil
Subject:	Final Performance Report to Dr. Tristan Nguyen
Contract/Grant Title:	Discovery of Empirical Components by Information Theory, Random Matrix Theory, and Computational Topology
Contract/Grant #:	FA9550-13-1-0076
Reporting Period:	15 August 2013 to 14 August 2016

Participants:

Amit Singer (PI): Professor of Mathematics, Princeton University

Robert Calderbank (PI): Professor of Electrical Engineering, Duke University

Ingrid Daubechies (PI): Professor of Mathematics, Duke University

Afonso Bandeira: Graduate Student, Princeton University

Tejal Bhamre: Graduate Student, Princeton University

Yutong Chen: Graduate Student, Princeton University

Joao Morais Carreira Pereira: Graduate Student, Princeton University

Colin Sandon: Graduate Student, Princeton University

Matthew Nokleby: Postdoctoral Research Fellow, Duke University

Andrew Thompson: Postdoctoral Research Fellow, Duke University

Alireza Vahid: Postdoctoral Research Fellow, Duke University

Juhwan Yoo: Postdoctoral Research Fellow, Duke University

Jiaji Huang: Graduate Student, Duke University

Rujie Yin: Graduate Student, Duke University

Tingran Gao: Graduate Student, Duke University

Jameson Cahill: Postdoctoral Research Fellow, Duke University

Principal Activities and Findings:

Spectrum of Random Kernel Matrices

We derive the limiting spectral density of random matrices whose (i, j)-th entry is f(X_i^T X_j), where X_1, ...,X_n are i.i.d. standard Gaussian random vectors in R^p, and f is a real-valued function. The eigenvalue distribution of these kernel random matrices is studied in the high dimensional / large sample regime ("large p, large n"). Our analysis applies as long as the rescaled kernel function is generic, and particularly, this includes non-smooth functions, e.g. Heaviside step function. Interestingly, the limiting densities interpolate between the Marcenko-Pastur density and the Wigner semi-circle density.

Documentation: J22

Robust Principal Component Analysis

We proved that for data generated from an elliptical distribution, the limiting distribution of Tyler's M-estimator for the covariance matrix converges to a Marcenko-Pastur-type distribution. Elliptical distributions play an important role in portfolio theory, radar, and financial data, and are typically used whenever the empirical distributions are heavy-tailed due to outliers.

Documentation: J12

Principal Component Analysis from Noisy Projected Data

The sample covariance is the most popular way to estimate the covariance matrix of a dataset. However, in many situations the sample covariance cannot be formed directly from the measurements. For example, when there is missing data or when the measurements are linear projections of the underlying signals. While it is possible to estimate the low rank structure through the matrix completion/sensing framework, solutions of the latter can be obtained using either semidefinite program (nuclear norm minimization) which is slow in practice or alternating minimization that lacks in theoretical guarantees. We show that the low rank structure can be estimated via a solution of a linear system that is formed using tools from high dimensional PCA and suitable eigenvalue shrinkage. We applied this new methodology for the denoising of extremely noise cryo-electron microscopy images and to reveal three-dimensional structural variability in such datasets.

Documentation: J11, J19, C31

Compressive Sensing - Random Demodulator:

The sampling rate of analog-to-digital converters is severely limited by underlying technological constraints. Recently, Tropp et al. proposed a new architecture, called a random demodulator, that attempts to overcome this limitation by sampling sparse, band limited signals at a rate much

lower than the Nyquist rate. An integral part of this architecture is a random bi-polar modulating waveform (MW) that changes polarity at the Nyquist rate of the input signal. Technological constraints also limit how fast such a waveform can change polarity, so we propose an extension of the random demodulator that uses a run-length limited (RLL) modulating waveform, and which we call a constrained random demodulator (CRD). The RLL modulating waveform changes polarity at a slower rate. We establish that a CRD enjoys theoretical guarantees similar to the RD and that these guarantees are directly related to the power spectrum of the MW. Further, we show that the relationship between the placement of energy in the spectrum of the input signal and the placement of energy in the power spectrum of the MW has a major effect on the reconstruction performance of signals sampled by a CRD.

Documentation: J1, J8, C11, C12, C16

Compressive Sensing – Information Theoretic Limits

We approach the problem of how to design optimal measurements through the Singular Value Decomposition or SVD. The SVD is the product of three matrices and each plays a role in the design of optimal linear measurements. The function of the right eigenvectors in the SVD of the measurement matrix □ is to collect energy from the source, so they should coincide with the eigenvectors of the source covariance. We arrange these eigenvectors in decreasing order of the corresponding singular vectors, starting with the biggest and going down. The function of the left eigenvectors in the SVD of the measurement matrix □ is to align high energy source modes with low noise modes, so they should coincide with the eigenvectors of the noise covariance. We arrange these eigenvectors in increasing order of the corresponding singular vectors, starting with the smallest and going up. Finally, the function of the singular values of the measurement matrix □ is to distribute the available energy among the channel modes. Note that we have ordered eigenvalues so that when we consider the ratio of the ith noise singular value to the ith source singular value these ratios are increasing.

We design measurement matrices to maximize mutual information I(x; y), because we think about using conditional mean estimation to recover the signal of interest from the noisy projection. The minimum mean squared error is the trace of the MMSE matrix, the lower bound on the MMSE is minimized when the mutual information I(x; y) is maximized, and the inequality in the lower bound is met with equality when x is Gaussian. Our work takes advantage of a relationship between the gradient of mutual information and the MMSE matrix that was discovered by Guo, Shamai and Verdú in 2005.

The work of Verdú and collaborators is motivated by communications, where the aim is to maximize mutual information between input signal and received signal. In communications we know the statistics of the source x, that is to say we know the correlation matrix □$_x$ and we can calculate its singular value decomposition. If we know the channel, and hence its SVD, then we can align the source so that it is minimally attenuated by the channel. That is the function of the precoder designed by Palomar and Verdu that is inserted between the transmitter and the channel. In sensing we know the SVD of the source and we are simply trying to design the SVD of the measurement matrix.

We have developed a generalization of Bregman divergence to unify vector Poisson and Gaussian channels. We are interested in vector Poisson channels because they are a good

model for X-ray scatter, also for document classification. In document classification we assume L classes of documents each characterized by a vector of probabilities over n words. The Poisson model describes how the words are drawn for a document in a given class. The number of words is large so we count the number of times words in subsets of the dictionary appear. These subsets act like key words associated with a given topic – these are our compressive measurements. We have applied our theory to compressive topic modeling for analysis of document corpora, and improves upon the state of the art for the 20 Newsgroups corpus.

Documentation: J5, C3, C10, C15, C17, C18, C20, C22, C26

Compressive Sensing – Subspace Models

Many important types of signal, including speech, faces, digits and fingerprints, can be accurately modeled as low-dimensional subspaces in a larger ambient space. Hence the problem of using a limited number of linear measurements to discriminate subspaces excited by Gaussian noise is fundamental to modern detection and estimation.

We are able to determine to within a single measurement the minimum number of measurements required to successfully reconstruct a signal drawn from a Gaussian mixture distribution in the low noise regime. Our method is to develop upper and lower bounds that are a function of the maximum dimension of the linear subspaces spanned by the Gaussian mixture components. We show that an n-dimensional signal that is s-sparse with non-zero components drawn independent identically distributed from a Gaussian mixture distribution can be reconstructed perfectly in the low-noise regime with exactly s+1 measurements. This estimate is tighter and sharper than standard bounds on the minimum number of measurements needed to recover sparse signals associated with a union of subspaces model. It shows that it is possible to achieve the performance of intractable l_0-pseudonorm recovery algorithms using the optimal closed-form conditional mean estimator within the Bayesian compressive sensing paradigm.

We derive these results by developing a first-order low-noise expansion of the MMSE that captures the existence or absence of an MMSE floor as well as the rate of decay to this floor. The presence or absence of an MMSE floor depends only on the relation between the number of measurements t and the rank s of the source covariance. The exact value of the MMSE floor (when t is less than s) and the MMSE power offset (when t is at least s) depends on the relation between the geometry of the measurement kernel and the geometry of the source. This geometric relation is captured by a multivariate generalization of the MMSE dimension (introduced by Wu and Verdu in 2011) that distinguishes MMSE expansions associated with different measurement kernels and source covariances. We are then able to use this geometric framework to quantify the advantage of measurement kernels that are designed over those that are random. While kernel design does not impact the phase transition, We are able to show that designed kernels can improve reconstruction performance both in terms of a lower error floor (if present) and a lower power offset. We have also connected theory to the practice of image reconstruction using a 20 class Gaussian mixture model for non-overlapping 8x8 image patches derived from 100,000 patches randomly sampled from 500 images in the Berkeley Segmentation Dataset. The phase transition phenomenon is clearly visible in our reconstruction of the image *Barbara* (which was not of course included in the original training set).

Documentation: J2, J4, J6, J7, J10, C2, C4, C5, C6, C8, C9, C22, C25

Deep Learning

Deep neural networks have proved very successful in domains where large training sets are available, but when the number of training samples is small, their performance suffers from overfitting. Prior methods of reducing over-fitting such as weight decay, Dropout and DropConnect are data-independent. Our work also motivated by the problem of overfitting, but the framework for learning features that are robust to data variation is different, and we are able to explicitly tradeoff the discriminative value of learned features against the generalization error of the learning algorithm. Our theoretical analysis starts with a cover of the data space, which is a partition into subsets with the property that distance between pairs in the same subset is bounded by □. We achieve robustness by encouraging the transform that maps data to features to be a local isometry, so that distances can increase by at most □. All that remains is to relate loss to distance, and we are able to achieve (K, 2A(□□□))-robustness, where A is the Lipschitz constant of the loss function.

Documentation: C19, C30, C32

Compressive Classification

We have shown that fundamental limits on classification cannot be avoided in a world where there is mismatch between a class and the subspace used to model that class. Our method is to connect the problem of using a limited number of linear measurements to discriminate subspaces, to that of using multiple transmit antennas to communicate over a non-coherent wireless channel. This connection, between two very different fields, means that capacity results obtained by Zheng and Tse for wireless communication can be used to derive fundamental limits on compressive classification. When a classifier tries to identify k-dimensional subspaces from an M-dimensional projection, corrupted by *noise/ mismatch* of variance $□^2$, we have shown that classification fails with high probability when there are more than $(1/□)^{M-k}$ subspaces to discern. When k is at least M/2 the converse holds true; classification succeeds with high probability when there are fewer than $(1/□)^{M-k}$ subspaces to discern.

Rate-distortion theory is the branch of information theory that deals with the lossy compression of random sources. Shannon's famous rate-distortion theorem relates the encoding rate R and the expected distortion according to the mutual information between the source and its estimate. Ahmad proposed to use rate-distortion analysis to bound learning performance, by treating the posterior distribution as a *soft* version of the MAP classifier. The posterior distribution is a random object, and it takes the role of the *source*, which we want to represent up to some distortion. The training samples take the role of the finite rate encoding of the posterior. The higher the number of samples the more information is conveyed about the posterior. The distortion measure is the average l_1 distance between the posterior and the estimate produced by the learning machine, and a classic result is that the generalization error is bounded above by the l_1 loss. We have used the machinery of rate-distortion theory to derive bounds on the tradeoff between classifier performance and the size of the training set. These bounds involve a quantity called the *Interpolation Dimension* that captures inherent complexity of the posterior. Interpolation dimension plays a role similar to the VC dimension in the classical theory, but provides bounds that are much tighter, particularly when the number of training samples is small.

Documentation: J9, C7, C14, C24, C27, C29

Wireless Communication

We have developed protocols that are able to take advantage of stale channel feedback. We have shown that if channel statistics are known, then it is possible to anticipate the statistics of collisions, and to transmit linear combinations of inputs that can be resolved at the receivers.

Documentation: C21, C28, C31

Data Storage

Use of Flash memory is increasing because capacity is increasing, and the cost differential between Flash and other storage technologies (especially hard drives) is narrowing. NAND Flash dominates solid-state drives (SSDs) and typical storage devices use multi-level cells with 2 (SLC), 4 (MLC) or 8 (TLC) levels per cell. MLCs are usually preferred because they are more mature than TLCs and provide better storage density than SLCs. One drawback to using Flash is that we can only erase a Flash cell a given number of times before that cell can no longer retain information. The number of Program/Erase (P/E) cycles that a cell can tolerate depends on the type of the cell used (SLC, MLC or TLC), and the scale of the Flash technology. Another practical difficulty is that the 4 physical levels per MLC cell are accessible only as two virtual 2-level cells on separate pages. We have developed a method of creating virtual Flash cells with several logical levels that avoids the need to change current hardware. We have demonstrated how to implement waterfall coding on the new virtual cells, and have introduced a new pseudo-erase operation that further extends memory lifetime. Our work connects the current Flash interface with the promise of coding techniques developed by the information theory and coding community.

Documentation: C1, C23

Publications:

Journal Papers:

1. H.A. Harms, W.U. Bajwa and R. Calderbank, A constrained random demodulator for sub-Nyquist sampling, IEEE Transactions on Signal Processing, Vol. 61 (3), pp. 707-723, February 2013
2. M.F. Duarte, S. Jafarpour and R. Calderbank, Performance of the Delsarte-Goethals frames on clustered sparse vectors, IEEE Transactions on Signal Processing, Vol. 61 (8), pp. 1998-2008, April 2013.
3. M. Nokleby, W.U. Bajwa, R. Calderbank and B. Aazhang, Toward Resource-Optimal Consensus over the Wireless Medium, IEEE Journal of Selected Topics in Signal Processing, Vol. 7 (2), pp. 284-295, April 2013.

4. Y. Chi, Y.C. Eldar, and R. Calderbank, PETRELS: Parallel Subspace Estimation and Tracking by Recursive Least Squares from Partial Observations, IEEE Transactions on Signal Processing, Vol. 61 (23), pp. 5947-5959, December 2013
5. L. Wang, D.E. Carlson, M.R.D. Rodrigues, R. Calderbank, and L. Carin, A Bregman matrix and the gradient of mutual information for vector Poisson and Gaussian channels, IEEE Transactions on Information Theory, Vol. 60 (5), pp. 2611-2629, May 2014
6. F. Renna, R. Calderbank, L. Carin, and M.R.D. Rodrigues, Reconstruction of signals drawn from a Gaussian mixture via noisy compressive measurements, IEEE Transactions on Signal Processing, Vol. 62 (9), pp. 2265-2277, September 2014
7. R. Calderbank, A. Thompson, and Y. Xie, On block coherence of frames, Journal of Applied and Computational Harmonic Analysis, Vol. 38 (1), pp. 50-71, January, 2015
8. H. A. Harms, W.U. Bajwa, and R. Calderbank, Identification of linear time-varying systems through waveform diversity, IEEE Transactions on Signal Processing, Vol. 63 (8), April 2015
9. M. Nokleby, M.R.D. Rodrigues, and R. Calderbank, Discrimination on the Grassmann Manifold: Fundamental limits of subspace classifiers, IEEE Transactions on Information Theory, Vol. 61 (4), pp. 2133-2147, April 2015
10. W.U. Bajwa, M.F. Duarte, and R. Calderbank, Conditioning of random block subdictionaries with applications to block-sparse recovery and regression, IEEE Transactions on Information Theory, Vol. 61 (7), pp. 4060-4079, July 2015
11. T. Bhamre, T. Zhang, A. Singer, Denoising and Covariance Estimation of Single Particle Cryo-EM Images, Journal of Structural Biology, accepted. Available at http://arxiv.org/abs/1602.06632.
12. T. Zhang, X. Cheng, A. Singer, Marchenko-Pastur Law for Tyler's and Maronna's M-estimators, Journal of Multivariate Analysis, accepted. Available at http://arxiv.org/abs/1401.3424.
13. A. Singer, H.-T. Wu, Spectral Convergence of the Connection Laplacian from Random Samples, Information and Inference: A Journal of the IMA, accepted. Available at http://arxiv.org/abs/1306.1587.
14. A. S. Bandeira, C. Kennedy, A. Singer, Approximating the Little Grothendieck Problem over the Orthogonal Group, Mathematical Programming, Series A, accepted. Available at http://arxiv.org/abs/1308.5207.
15. Z. Zhao, Y. Shkolnisky, A. Singer, Fast Steerable Principal Component Analysis, IEEE Transactions on Computational Imaging, 2 (1), pp. 1-12, 2016.
16. O. Özyeşil, A. Singer, R. Basri, Stable Camera Motion Estimation using Convex Programming, SIAM Journal on Imaging Sciences, 8 (2), pp. 1220-1262, 2015.
17. C. J. Dsilva, B. Lim, H. Lu, A. Singer, I. G. Kevrekidis, S. Y. Shvartsman, Temporal ordering and registration of images in studies of developmental dynamics, Development, 142, 1717-1724, 2015.
18. K. N. Chaudhury, Y. Khoo, A. Singer, Global registration of multiple point clouds using semidefinite programming, SIAM Journal on Optimization, 25 (1), pp. 468-501, 2015.
19. E. Katsevich, A. Katsevich, A. Singer, Covariance Matrix Estimation for the Cryo-EM Heterogeneity Problem, SIAM Journal on Imaging Sciences, 8 (1), pp. 126-185, 2015.

20. E. Abbe, A. S. Bandeira, A. Bracher, A. Singer, Decoding binary node labels from censored edge measurements: Phase transition and efficient recovery, IEEE Transactions on Network Science and Engineering, 1 (1) pp. 10-22, 2014.
21. Z. Zhao, A. Singer, Rotationally Invariant Image Representation for Viewing Direction Classification in Cryo-EM, Journal of Structural Biology, 186 (1), pp. 153-166, 2014.
22. X. Cheng, A. Singer, The Spectrum of Random Inner-Product Kernel Matrices, Random Matrices: Theory and Applications, 2 (4) 1350010 (47 pages), 2013.

Conference Papers:

1. A.N. Jacobvitz, R. Calderbank and D.J. Sorin, Coset coding to extend the lifetime of memory, Proceedings of the 19th IEEE International Symposium on High Performance Computer Architecture (HPCA 2013), pp. 222-233, Shenzhen, China, February 2013
2. M. Wang, W. Xu and R. Calderbank, Compressed sensing with corrupted participants, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 4653-4657, Vancouver, Canada, May 2013
3. F. Renna, M.R.D. Rodrigues, M. Chen, R. Calderbank and L. Carin, Compressive sensing for incoherent imaging systems with optical constraints, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 5484-5488, Vancouver, Canada, May 2013
4. Y. Chi and R. Calderbank, Knowledge-enhanced matching pursuit, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 6576-6580, Vancouver, Canada, May 2013
5. H. Reboredo, F. Renna, R. Calderbank and M.R.D Rodrigues, Compressive classification, Proceedings of the IEEE International Symposium on Information Theory, pp. 1616-1620, Istanbul, Turkey, July 2013
6. T. Wu, G. Polatkan, D. Steel, W. Brown, I. Daubechies and R. Calderbank, Painting analysis using wavelets and probabilistic topic models, Proceedings of the IEEE International Conference on Image Processing (ICIP 2013), Melbourne, Australia, September 2013.
7. M. Nokleby, R. Calderbank and M.R.D. Rodrigues, Information-theoretic limits on the classification of Gaussian mixtures: Classification on the Grassmann manifold, Proceedings of the IEEE Information Theory Workshop (ITW 2013), Seville, Spain, September 2013
8. H. Robredo, F. Renna, R. Calderbank and M.R.D. Rodrigues, Projection designs for compressive classification, 1st IEEE Global Conference on Signal and Image Processing, Austin, Texas, December 2013
9. F. Renna, R. Calderbank, L. Carin and M.R.D. Rodrigues, Reconstruction of Gaussian mixture models from compressive measurements: A phase transition view, 1st IEEE Global Conference on Signal and Image Processing, Austin, Texas, December 2013
10. L. Wang, D.E. Carlson, M.R.D Rodrigues, D. Wilcox, R. Calderbank and L. Carin, Designed measurements for vector count data, Advances in Neural Information Processing Systems 26 (NIPS), Lake Tahoe, Nevada, December 2013

11. H. A. Harms, W.U. Bajwa, and R. Calderbank, Resource-efficient parametric recovery of linear time-varying systems, Proceedings of the 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2013), Saint Martin, December 2013
12. H. A. Harms, W.U. Bajwa, and R. Calderbank, Shaping the power spectra of bipolar sequences with application to sub-Nyquist sampling, Proceedings of the 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2013), Saint Martin, December 2013
13. M. Nokleby, M.R.D. Rodrigues, and R. Calderbank, Information-theoretic criteria for the design of compressive subspace classifiers, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 3067-3071, Florence, Italy, May 2014
14. M. Nokleby, M.R.D. Rodrigues, and R. Calderbank, Discrimination on the Grassmann manifold: Fundamental limits of subspace classifiers, Proceedings of the IEEE International Symposium on Information Theory, pp. 3012-3016, Honolulu, Hawaii, June 2014
15. L. Wang, A. Razi, M.R.D. Rodrigues, R. Calderbank, and L. Carin. Nonlinear Information-Theoretic Compressive Measurement Design, Proceedings of the International Conference on Machine Learning (ICML), Beijing, China, June 2014.
16. H. A. Harms, W.U. Bajwa, and R. Calderbank, Efficient linear time-varying system Identification using chirp waveforms, Proceedings of the Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, California, November 2014
17. J. Huang, X. Yuan, and R. Calderbank, Multi-scale Bayesian reconstruction of compressive x-ray image, Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2015). Brisbane, Australia, April 2015
18. J. Huang, X. Yuan, and R. Calderbank, Collaborative compressive x-ray image reconstruction, Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2015). Brisbane, Australia, April 2015
19. J. Huang, Q. Qiu, R. Calderbank, M.R.D. Rodrigues, and G. Sapiro, Alignment with intra-class structure can improve classification, Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2015). Brisbane, Australia, April 2015
20. X. Yuan, J. Huang, and R. Calderbank, Polynomial-phase signal direction-finding and source-tracking with a single acoustic vector sensor, Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2015). Brisbane, Australia, April 2015
21. A. Vahid and R. Calderbank, Impact of local delayed CSIT on the capacity region of the two-user interference channel, Proceedings of the IEEE International Symposium on Information Theory, Hong Kong, China, June 2015
22. F. Renna, L. Wang, X. Yuan, J. Yang, G. Reeves, R. Calderbank, L. Carin and M.R.D. Rodrigues, Classification and reconstruction of compressed GMM signals with side information, Proceedings of the IEEE International Symposium on Information Theory, Hong Kong, China, June 2015

23. I. Tamo, A. Barg, S. Goparaju, R. Calderbank, Cyclic LRC codes and their subfield subcodes, Proceedings of the IEEE International Symposium on Information Theory, Hong Kong, China, June 2015
24. A. Beirami, R. Calderbank, K. Duffy, and M. Medard, Computational security subject to source constraints, guesswork and inscrutability, Proceedings of the IEEE International Symposium on Information Theory, Hong Kong, China, June 2015
25. J. Sokolic, F. Renna, R. Calderbank, and M.R.D. Rodrigues, Mismatch in the classification of linear subspaces: Upper bound to the probability of error, Proceedings of the IEEE International Symposium on Information Theory, Hong Kong, China, June 2015
26. L. Wang, J. Huang, X. Yuan, V. Cevher, M.R.D Rodrigues, R. Calderbank, and L. Carin, A concentration-of-measure inequality for multiple-measurement models, Proceedings of the IEEE International Symposium on Information Theory, Hong Kong, China, June 2015
27. A. Beirami, R. Calderbank, M. Christiansen, K. Duffy, A. Makhdoumi, and M. Medard, A geometric perspective on guesswork, to appear in Proceedings of the 53rd Annual Allerton Conference on Communication, Control, and Computing, Monticello, Illinois, September 2015
28. A. Vahid, I. Shomorony, and R. Calderbank, Informational bottlenecks in two-unicast wireless networks with delayed CSIT, Proceedings of the 53rd Annual. Allerton Conference on Communication, Control, and Computing, Monticello, Illinois, September 2015
29. M. Nokleby, A. Beirami, and R. Calderbank, A rate-distortion framework for supervised learning, Proceedings of IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2015), Boston, Massachusetts, September, 2015
30. O. Özyeşil, A. Singer, Robust Camera Location Estimation by Convex Programming, in Computer Vision and Pattern Recognition (CVPR 2015), pp. 2674-2683, 7-12 June 2015.
31. J. Andén, E. Katsevich, A. Singer, Covariance estimation using conjugate gradient for 3D classification in Cryo-EM, in IEEE 12th International Symposium on Biomedical Imaging (ISBI 2015), pp. 200-204, 16-19 April 2015.
32. T. Bhamre, T. Zhang, A. Singer, Orthogonal Matrix Retrieval in Cryo-Electron Microscopy, in IEEE 12th International Symposium on Biomedical Imaging (ISBI 2015), pp. 1048-1052, 16-19 April 2015.
33. K. N. Chaudhury, Y. Khoo, A. Singer, Large-scale sensor network localization via rigid subnetwork registration, in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2015), pp.2849--2853, 19-24 April 2015.
34. A. S. Bandeira, Y. Khoo, A. Singer, Open problem: Tightness of maximum likelihood semidefinite relaxations, in 2014 Conference on Learning Theory (COLT 2014), JMLR: Workshop and Conference Proceedings vol 35:1–3, 2014.
35. E. Abbe, A. S. Bandeira, A. Bracher, and A. Singer, Linear Inverse problems on Erdős-Rényi graphs: Information-theoretic limits and efficient recovery, in IEEE International Symposium on Information Theory (ISIT 2014), pp. 1251-1255, June 29-July 4 2014.
36. A. S. Bandeira, M. Charikar, A. Singer, A. Zhu, Multireference Alignment using Semidefinite Programming, in Proceedings of the 5th conference on Innovations in Theoretical Computer Science (ITCS '14), pp. 459-470

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
05-12-2016

2. REPORT TYPE
Final Performance Report

3. DATES COVERED *(From - To)*
02/15/2013 – 02/14/2016

4. TITLE AND SUBTITLE
Discovery of Empirical Components by Information Theory, Random Matrix Theory, and Computational Topology

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-13-1-0076

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Singer, Amit; Calderbank, Robert; Daubechies, Ingrid

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Princeton University
Department of Mathematics and The
Program in Applied and Computational
Mathematics
202 Fine Hall
Princeton, NJ 08544

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
875 North Randolph Street, Room 3112
Arlington, VA 22203-1768

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
All outcomes are publicly available.

13. SUPPLEMENTARY NOTES
N/A

14. ABSTRACT
Recent years have seen exciting new developments in mathematics and computer science, which have opened up new domains of application for computational mathematics. These come with new challenges, for which new approaches and tools must be and are being developed. Machine learning and compressive sensing are two typical examples; they draw not only from traditional linear algebra based numerical analysis or approximation theory, but also from information theory, graph theory, the geometry of Banach spaces, probability theory, and more.
This proposal seeks to fund research of three faculty drawn to these new computational challenges, who are also finding increasingly that their different fields of expertise all contributes to the development of dramatically more effective tools. The confluence of interests, and the conviction that joining their efforts will produce a whole that exceeds the sum of its parts, constitute the engine that drives the approaches proposed here.

15. SUBJECT TERMS
Compressive Sensing, Subspace Modelling, Information Theory, Random Matrix Theory

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Lisa D. Giblin
U	U	U	UU		**19b. TELEPHONE NUMBER** *(include area code)* (609) 258-5128

FEDERAL FINANCIAL REPORT

(Follow form instructions)

1. Federal Agency and Organizational Element to Which Report is Submitted Air Force Office of Scientific Research 875 North Randolf Street Suite 325, Room 3112 Arlington VA 22203	2. Federal Grant or Other Identifying Number Assigned by Federal Agency (To report multiple grants, use FFR Attachment) FA9550-13-1-0076	Page 1 of 1 pages
3. Recipient Organization (Name and complete address including Zip code) The Trustees of Princeton University 701 Carnegie Center, Suite 443 Princeton, NJ 08540		

4a. DUNS Number	4b. EIN	5. Recipient Account Number or Identifying Number (To report multiple grants, use FFR Attachment)	6. Report Type	7. Basis of Accounting
002484665	21-0634501	CNV1002205	☐ Quarterly ☐ Semi-Annual ☐ Annual ☑ Final	☑ Cash ☐ Accrual

8. Project/Grant Period (Month, Day, Year)		9. Reporting Period End Date (Month, Day, Year)
From: 02/15/2013	To: 02/14/2016	02/14/2016

10. **Transactions**	Cumulative
(Use lines a-c for single or combined multiple grant reporting)	
Federal Cash (To report multiple grants separately, also use FFR Attachment):	
a. Cash Receipts	723,249.38
b. Cash Disbursements	889,344.85
c. Cash on Hand (line a minus b)	(166,095.47)
(Use lines d-o for single grant reporting)	
Federal Expenditures and Unobligated Balance:	
d. Total Federal funds authorized	900,000.00
e. Federal share of expenditures	889,344.85
f. Federal share of unliquidated obligations	-
g. Total Federal share (sum of lines e and f)	889,344.85
h. Unobligated balance of Federal funds (line d minus g)	10,655.15
Recipient Share:	
i. Total recipient share required	-
j. Recipient share of expenditures	-
k. Remaining recipient share to be provided (line i minus j)	-
Program Income:	
l. Total Federal share of program income earned	
m. Program income expended in accordance with the deduction alternative	
n. Program income expended in accordance with the addition alternative	
o. Unexpended program income (line l minus line m or line n)	

11. Indirect Expense	a. Type	b. Rate	c. Period From	Period To	d. Base	e. Amount Charged	f. Federal Share
	FIXED	61%	2/15/2013	2/14/2016	168,754.06	102,940.02	102,940.00
				g. Totals:	168754.06	102,940.02	102,940.00

12. Remarks: Attach any explanations deemed necessary or information required by Federal sponsoring agency in compliance with governing legislation:

13. Certification: By signing this report, I certify to the best of my knowledge and belief that the report is true, complete, and accurate, and the expenditures, disbursements and cash receipts are for the purposes and intent set forth in the award documents. I am aware that any false, fictitious, or fraudulent information may subject me to criminal, civil, or administrative penalties. (U.S. Code, Title 18, Section 1001)

a. Typed or Printed Name and Title of Authorized Certifying Official Alberta Molnar Manager, Sponsored Research Accounting	c. Telephone (Area code, number, and extension) 609-258-3070 d. Email Address sra@princeton.edu
b. Signature of Authorized Certifying Official Alberta Molnar	e. Date Report Submitted (Month, Day, Year) 5/10/2016
	14. Agency use only:

Standard Form 425 - Revised 10/11/2011
OMB Approval Number: 0348-0061
Expiration Date: 2/28/2015

1.

1. Report Type

Final Report

Primary Contact E-mail
Contact email if there is a problem with the report.

LGiblin@princeton.edu

Primary Contact Phone Number
Contact phone number if there is a problem with the report

609-258-5128

Organization / Institution name

Princeton University

Grant/Contract Title
The full title of the funded effort.

Discovery of Empirical Components by Information Theory, Random Matrix Theory, and Computational Topology

Grant/Contract Number
AFOSR assigned control number. It must begin with "FA9550" or "F49620" or "FA2386".

FA9550-13-1-0076

Principal Investigator Name
The full name of the principal investigator on the grant or contract.

Dr. Amit Singer

Program Manager
The AFOSR Program Manager currently assigned to the award

Dr. Tristan Nguyen

Reporting Period Start Date

02/15/2013

Reporting Period End Date

02/14/2016

Abstract

Recent years have seen exciting new developments in mathematics and computer science, which have opened up new domains of application for computational mathematics. These come with new challenges, for which new approaches and tools must be and are being developed. Machine learning and compressive sensing are two typical examples; they draw not only from traditional linear algebra based numerical analysis
or approximation theory, but also from information theory, graph theory, the geometry of Banach spaces, probability theory, and more.
This proposal seeks to fund the research of three faculty drawn to these new computational challenges, who are also finding increasingly that their different fields of expertise all contribute to the development of dramatically more effective tools. This confluence of interests, and the conviction that joining their efforts will produce a whole that exceeds the sum of its parts, constitute the engine that drives the approaches proposed here.

Distribution Statement
This is block 12 on the SF298 form.

Distribution A - Approved for Public Release

Explanation for Distribution Statement

If this is not approved for public release, please provide a short explanation. E.g., contains proprietary information.

SF298 Form

Please attach your SF298 form. A blank SF298 can be found here. Please do not password protect or secure the PDF The maximum file size for an SF298 is 50MB.

Standard Form 298, Report Documentation Page.pdf

Upload the Report Document. File must be a PDF. Please do not password protect or secure the PDF . The maximum file size for the Report Document is 50MB.

Final Performance Report Electronic Submission.pdf

Upload a Report Document, if any. The maximum file size for the Report Document is 50MB.

Archival Publications (published) during reporting period:

Changes in research objectives (if any):

None

Change in AFOSR Program Manager, if any:

Yes, previously AFOSR Program Manager was Dr. Robert Bonneau

Extensions granted or milestones slipped, if any:

None

AFOSR LRIR Number

LRIR Title

Reporting Period

Laboratory Task Manager

Program Officer

Research Objectives

Technical Summary

Funding Summary by Cost Category (by FY, $K)

	Starting FY	FY+1	FY+2
Salary			
Equipment/Facilities			
Supplies			
Total			

Report Document

Report Document - Text Analysis

Report Document - Text Analysis

Appendix Documents

2. Thank You

E-mail user

May 12, 2016 16:17:24 Success: Email Sent to: LGiblin@princeton.edu